Adaptive soil moisture profile filtering for horizontal information propagation in the independent column-based CLM2.0

Gabriëlle J.M. De Lannoy *

Laboratory of Hydrology and Water Management, Ghent University, Coupure links 653, B-9000 Ghent, Belgium,

George Mason University & Center for Research on Environment and Water, 4041 Powder Mill Road, Suite 302, Calverton, MD 20705-3106, USA

Paul R. Houser

George Mason University & Center for Research on Environment and Water, 4041 Powder Mill Road, Suite 302, Calverton, MD 20705-3106, USA

Niko E.C. Verhoest

Laboratory of Hydrology and Water Management, Ghent University, Coupure links 653, B-9000 Ghent, Belgium

Valentijn R.N. Pauwels

Laboratory of Hydrology and Water Management, Ghent University, Coupure links 653, B-9000 Ghent, Belgium

* *Corresponding author address:* Gabriëlle J.M. De Lannoy, Laboratory of Hydrology and Water Management, Ghent University, Coupure links 653, B-9000 Ghent, Belgium, and George Mason University & Center for Research on Environment and Water, 4041 Powder Mill Road, Suite 302, Calverton, MD 20705-3106, USA
E-mail: Gabrielle.DeLannoy@UGent.be

ABSTRACT

Data assimilation aims to provide an optimal estimate of the overall system state, not only for an observed state variable or location. However, large scale land surface models are typically column-based and purely random ensemble perturbation of states will lead to block-diagonal a priori (or background) error covariances. This facilitates the filtering calculations, but compromises the potential of data assimilation to influence (unobserved) vertical and horizontal neighboring state variables. Here, a combination of an ensemble Kalman filter and an adaptive covariance correction method is explored to optimize the variances and retrieve the off-block-diagonal correlations in the a priori error covariance matrix. In a first time period, all available soil moisture profile observations in a small agricultural field are assimilated into the CLM2.0 land surface model to find the adaptive second order a priori error information. After that period, only observations from single individual soil profiles are assimilated with inclusion of this adaptive information. It is shown that assimilation of a single profile can partially rectify the incorrectly simulated soil moisture spatial mean and variability. The largest reduction in the root spatial mean square error in the soil moisture field varies between 7 and 22%, depending on the soil depth, when assimilating a single complete profile every 2 days during 3 months with a single time-invariant covariance correction.

Key words: adaptive filter, Kalman filter, soil moisture, ensemble, error covariance, spatial structure

1. Introduction

For many hydrological applications, it is desirable to update an entire spatial field through the assimilation of sparse soil moisture observations. Information propagation over a soil volume can be accomplished by full 3-dimensional (3D) Richards equation-based models. However, for computational efficiency, coarse scale land surface models (LSM) often treat soil moisture profiles as independent individual vertical columns, ignoring horizontal water flows. In a synthetic data assimilation study, Reichle and Koster (2003) demonstrated that taking horizontal model error correlations into account in the a priori error covariance matrix for a 3D ensemble Kalman filter improved the estimation accuracy for soil moisture, even in a regional scale application. Other existing approaches to enforce lateral information flow in an assimilation scheme with a 1D (column) model are a priori interpolation of the observations to locations where they are desired to affect the model state, as in analysis nudging (Stauffer and Seaman 1990; Houser et al. 1998), and combined spatial interpolation and temporal propagation as explored in the space-time Kalman filter (Huang and Cressie 1996). Generally, the choice is between a better or more complex LSM with a simple assimilation scheme or a simple LSM with a more complex assimilation scheme. Exploring the possibility of accounting for horizontal information flow within the assimilation scheme is worthwhile, because '*it is unreasonable to rely on models that relate each pixel of a large grid to every other pixel*' (McLaughlin 2002). It should also be recognized that for very coarse scale modelling, it is not needed to horizontally connect all the state variables within the model, because of the limited correlation length of earth state variables.

For several decades, it has been discussed that, for a given system model and observation

network (Daley 1992b), the efficient use of the information brought by the observations is largely determined by the second order a priori error statistics: the a priori (or background) error covariance matrix $\mathbf{P}_i^-$, used in the Kalman gain (or any blending matrix used during assimilation) at each time step i, distributes information in space and between state variables and is the key to optimal data assimilation. Furthermore, the magnitude of the blending matrix elements essentially only depend on the signal-to-noise ratio $\mathbf{R}_i/\mathbf{P}_i^-$, with $\mathbf{R}_i$ the observation error covariance matrix. Different blending gains mean different filter bandwidths and a distinct sensitivity to noise: the larger the a priori state error (or the smaller the observation error), the larger the filter bandwith and the more variation (due to incorporation of observations) can be expected in the assimilation analyses.

In the sequential statistical interpolation method (Daley 1991) and in most variational assimilation schemes, $\mathbf{P}_i^-$ is generally set empirically, after splitting $\mathbf{P}_i^-$ into a time-invariant correlation matrix $\mathbf{P_C}$ and a possibly time-variant diagonal matrix $\mathbf{P_{S}}_i$ whose elements are the standard deviations for the individual variables in the state, i.e. $\mathbf{P}_i^- = (\mathbf{P_{S}}_i\mathbf{P_C}\mathbf{P_{S}}_i)^-$. The specification of the correlation structure is a critical issue, because it selects the area from which observations are allowed to influence the model variables. Mostly, stationary, homogeneous, horizontal/vertical separable and isotropic error covariance functions as function of distance have been proposed for which only a few parameters were to be estimated, e.g. by the traditional 'innovation method' or observational method (Hollingsworth and Lönnberg 1986; Daley 1991; Houser et al. 1998). With the implementation of the Kalman filter, the focus was moved to the definition of the model error covariance (Daley 1992a), because $\mathbf{P}_i^-$ can easily be calculated by the explicit Lyapunov propagation equation in linear stationary systems as the sum of the predictability and model error covariances. For optimization, several

adaptive filters have been developed to estimate the model error covariances, mostly involving some whitening of the innovation (observation minus forecast) sequence (Kailath 1968; Mehra 1970; Maybeck 1982), or matching (elements of) the zero lag innovation covariance matrix to the theoretically optimal values (Jazwinski 1969; Myers and Tapley 1976; Maybeck 1982; Mehra 1972; Dee 1995; Dee and da Silva 1999). The majority of the techniques focuss more on the estimation of the error variances than on the error correlation. In non-linear systems, as most land surface models, ensembles allow a flow-dependent statistical approximation of $\mathbf{P}_i^-$ for the ensemble Kalman filter (EnKF, Evensen (1994)). However, the EnKF only quantifies the uncertainty in the space spanned by the ensembles and it is particularly difficult to reflect the 'true' uncertainty in the initial or analysis state and model error into an optimal perturbation magnitude and structure for the forcings, model structure, state and parameters. It is often useful to tune the filter by adapting the error variance, e.g. by inflation, to avoid model divergence in case of underestimated a priori error covariance. An adaptive inflation factor in time was proposed by Anderson (2007), who adjusted the variance of each state component, while the correlation between pairs of components remained unchanged. Mitchell and Houtekamer (2000) applied the maximum-likelihood method to estimate model error variance parameters in an EnKF framework, where initially only the uncertainty in the initial state or analysis was considered.

In hydrology, the statistical nature of the a priori (forecast) errors has often been prescribed in synthetical studies or arbitrarily chosen in real studies. Sometimes, the filter performance is evaluated afterwards by providing innovation statistics (Reichle et al. 2002a; Crow 2003; De Lannoy et al. 2007b). In synthetic studies, filters can be calibrated by contrasting the filter output to some 'truth', as in Reichle et al. (2002b) and Reichle and Koster

(2003) to allow a comparison of several Kalman filter setups for soil moisture estimation, each in optimal conditions. If no truth is available, the innovation statistics could be used to optimize the filter. van Geer et al. (1991) and Zhang et al. (2004) applied the covariance matching technique (Mehra 1972) to optimize the uncertainty estimation in groundwater and soil moisture simulations, respectively. In the covariance matching method, the innovation covariance matrix, obtained as a result from the filter, is iteratively contrasted to its theoretical value, which is a function of the tunable model error covariance matrix. The innovation covariance matrix is typically calculated as the outer product of the centralized innovation vectors over some time (proxy of ensemble). Drécourt et al. (2006) extended and changed this approach to estimate both the system and bias model error covariance in a synthetical groundwater simulation study. Crow and van Loon (2006) applied the maximum-likelihood estimation (Dee 1995) to find a covariance scaling parameter to adjust the error variances only, while assuming a fixed, predefined (spatial) correlation structure. This approach involves the inner product (or, the trace of the outer product) of the innovation vectors, i.e. all variables' uncertainties are lumped into 1 value. They illustrated how tuning the wrong error source to produce better innovation statistics led to a reduced accuracy of the filter results, when only assimilating surface soil moisture. Additionally, there was a risk to find the 'globally' best innovation statistics through optimization of the wrong error type, if different sources of error were taken into account. Reichle et al. (2008) used the same maximum-likelihood approach to identify model and observation error variances for soil moisture assimilation in a synthetic experiment. The multi-scale Kalman filter (Parada and Liang 2004; Kumar 1999) efficiently considers the spatial dependence and scaling properties of soil moisture. Parada and Liang (2004) applied a multi-scale Kalman filter in conjunction with an expectation

maximization algorithm to estimate temporally evolving statistical observation and model error parameters inherent to the Kalman filter to assimilate surface soil moisture in an LSM. Besides interest in the $\mathbf{P}_i^-$, also the observation uncertainty (error variance) is an important factor in filtering. In several observation system simulation experiments (OSSE), as e.g. by Walker and Houser (2004), one has tried to quantify the maximal acceptable (to be useful in data assimilation) observation uncertainty to guide instrumentation design. It is reasonable to assume that, similar to forecast error, observation error variance is also state dependent and hence an adaptive filtering approach could be beneficial (Romanowicz et al. 2006).

Here, in situ soil moisture profiles in a small agricultural field (Optimizing Production Inputs for Economic and Environmental Enhancement, OPE^3) are simulated by individual, unconnected, soil columns in the Community Land Model (CLM2.0). Ensembles are used to quantify the variables' uncertainty in each individual profile. Then, the ensemble a priori error variance is adapted and the error correlation between the profiles is sought for in an adaptive filtering scheme to introduce lateral flow of information through the assimilation scheme. The adaptive information is extracted from as many field observations as available. In a successive time period, this information is used to update the complete field while assimilating only a few observations. This study differs from the EnKF approach of Reichle and Koster (2003), who imposed and calibrated (against some truth) the model error field in a synthetic study to impose lateral flow of information. While data assimilation updates the observable prognostic state variables, it also adjusts the unobserved diagnostic flow variables. This study only focuses on the possibility of propagating soil moisture profiles horizontally and only validates observable prognostic soil moisture. An earlier study (De Lannoy et al. 2007b) highlighted the effect of soil moisture assimilation on diagnostic evapotranspiration,

subsurface drainage and runoff.

2. Model and observations

Soil moisture profiles are estimated with the Community Land Model (CLM2.0, Dai et al. (2003)) at 36 locations (Figure 1) in a 21 ha corn field in Beltsville, Maryland (USA), where the Optimizing Production inputs for Economic and Environmental Enhancement (OPE^3) experiment is conducted (Gish et al. 2002; De Lannoy et al. 2006b). The model simulates land surface processes by calculating hourly vertical water and heat fluxes and states for each profile without interaction between profiles, i.e. the linearized system matrix is block-diagonal. To limit the computational costs, only the 10-layer soil moisture prognostic variables are included in the assimilation scheme, i.e. the state vector $\hat{\mathbf{x}}_i$ has a dimension of 360. The depths of the soil nodes are set to 2.5, 5, 10, 20, 30, 50, 80, 120, 150 and 180 cm. The nonlinear land surface model, $\mathbf{f}_{i,i-1}$ propagates the state vector from time step $i-1$ to i:

$$\hat{\mathbf{x}}_i^- = \mathbf{f}_{i,i-1}(\hat{\mathbf{x}}_{i-1}^+, \mathbf{u}_i), \tag{1}$$

with $\mathbf{u}_i$ the meteorological forcings, which are hourly observed data and assumed spatially homogeneous over all profiles in this study. The initial condition for the land surface model is the a posteriori state estimate (or analysis) $\hat{\mathbf{x}}_{i-1}^+$. This a posteriori state estimate is obtained from updating the a priori state estimate $\hat{\mathbf{x}}_{i-1}^-$ at time $i-1$ (discussed below).

During our observation period from 1 May 2001 through 30 April 2002, hourly measurements were obtained from 36 capacitance probes (EnviroSCAN, SENTEK Pty Ltd., South

Australia), measuring soil moisture at 3 to 7 layers (H-probes: 10, 30 and 80 cm, M-probes: 10, 30, 50, 120, 150 and 180 cm, L-probes: 10, 30, 50, 80, 120, 150 and 180 cm depth, H-, M- and L-probes were installed at locations with expected different infiltration regimes, i.e. high, medium or low clay content.). A layout of the observation probes in the OPE^3 field is shown in Figure 1. The observations $\mathbf{y}_i$ (vector dimension varies with the specific assimilation experiment) are linearly related to the state by the observation (or forward) operator $\mathbf{H}_i$, i.e.

$$\mathbf{y}_i = \mathbf{H}_i \hat{\mathbf{x}}_i^- + \mathbf{v}_i, \tag{2}$$

with $\mathbf{v}_i$ zero mean assumed random error. $\mathbf{H}_i$ contains only values of 1 and 0, because the observations are direct measurements of the state variables.

The CLM was calibrated and initialized for each individual profile through weak constraint variational assimilation of observed data from 1 May 2001 through 1 October 2001 (De Lannoy et al. 2006a), with focus on parameter estimation in the period from 2 September 2001 through 2 October 2001. All state updating experiments in this paper start after this calibration period. A limited subset of observations will be assimilated for state and uncertainty estimation, while withholding other soil moisture profile observations for validation of the filtering performance. This study only focuses on assimilation of 'complete' profiles: depending on the type of probe and the availability of data, this means simultaneous assimilation of observations at 3 up to 7 soil depths.

3. Ensemble Kalman filter (EnKF) implementation

The Kalman filter is a sequential assimilation method to update forecasted model states with observational information. The analysis or a posteriori state estimate is a combination of the forecast and observations, each weighted as function of their uncertainty, which is captured in the a priori and observation error covariance matrix, respectively. The uncertainty in the forecasts is approximated by $N = 64$ ensemble (Monte Carlo) member predictions $\hat{\mathbf{x}}^-_{j,i}$ $(j = 1, \cdots, N)$:

$$\hat{\mathbf{x}}^-_{j,i} = \mathbf{f}_{i,i-1}(\hat{\mathbf{x}}^+_{j,i-1}, \mathbf{u}_i, \mathbf{w}_{j,i-1}), \tag{3}$$

with $\mathbf{w}_{j,i}$ a realization of the (zero mean assumed) model error that represents the complete effect of perturbations to forcings, parameters and state variables, quantified as in De Lannoy et al. (2006a). The a priori state (background) error covariance $\mathbf{P}^-_i$ is initially obtained from the ensemble sample covariance ($\mathbf{P}^-_i = \mathbf{P}^-_{ens,i}$). Due to the model structure (independent profiles) and our choice of uncorrelated forcing perturbations, a priori state errors at different locations were effectively uncorrelated. The spurious elements (without actual statistical significance) ending up outside of the diagonal blocks corresponding to the variables in one profile in $\mathbf{P}^-_{ens,i}$ were set to 0. This type of $\mathbf{P}^-_{ens,i}$ localization would be well suited for global or regional scale studies of variables that are only correlated over a small spatial domain and for which the error correlations over larger distances can be assumed to be very small.

When observations $\mathbf{y}_i$ are available, each ensemble member j is updated individually to obtain the a posteriori state estimate:

$$\hat{\mathbf{x}}^{+}_{j,i} = \hat{\mathbf{x}}^{-}_{j,i} + \mathbf{K}_i[\mathbf{y}_{j,i} - \mathbf{H}_i\hat{\mathbf{x}}^{-}_{j,i}]. \quad (4)$$

The actual analysis state $\hat{\mathbf{x}}^{+}_{i}$ is the ensemble mean. The term $[\mathbf{y}_{j,i} - \mathbf{H}_i\hat{\mathbf{x}}^{-}_{j,i}]$ equals the innovation vector $\mathbf{r}^{-}_{i}$. The observations are perturbed to ensure sufficient spread, i.e. $\mathbf{y}_{j,i} = \mathbf{y}_i + \mathbf{v}_{j,i}$, with $\mathbf{v}_{j,i}$ an imposed zero-mean Gaussian perturbation with an error covariance of $\mathbf{R} = (0.022\ \mathrm{m^3m^{-3}})^2 \cdot \mathbf{I}$, based on the sensor calibration information and assuming zero cross-correlation between the observation errors (De Lannoy et al. 2007b). The Kalman gain $\mathbf{K}_i$ is identical for all ensemble members and given by:

$$\mathbf{K}_i = \mathbf{P}^{-}_{i}\mathbf{H}^{T}_{i}[\mathbf{H}_i\mathbf{P}^{-}_{i}\mathbf{H}^{T}_{i} + \mathbf{R}_i]^{-1}. \quad (5)$$

Basically, the term $\mathbf{P}^{-}_{i}\mathbf{H}^{T}_{i}$ maps the analysis increments on the model grid. If no observations are available, then $\hat{\mathbf{x}}^{+}_{i}$ in Eq. 1 simply equals $\hat{\mathbf{x}}^{-}_{i}$.

4. Adaptive EnKF (ADEnKF)

a. Method

Within the blocks of $\mathbf{P}^{-}_{ens,i}$, one can assume that ensemble error correlations are well estimated: an assimilated observation in a single layer would be efficiently propagated to other layers within the same profile (Figure 2). However, the magnitude of the error variances could possibly be optimized, because it is defined by a subjectively chosen 'realistic' perturbation magnitude. Tuning the variances is typically the goal of most adaptive filtering schemes. Therefore, we will allow a variance correction within the adaptive scheme, while keeping the error correlations in the blocks as retrieved from the ensembles. Because the

ensemble approach does not allow the estimation of the error structure outside the diagonal blocks in $\mathbf{P}_i^-$, the adaptive scheme is also explored to retrieve this 'missing' between-profiles correlation information and hence allow horizontal information propagation (Figure 2).

Most adaptive schemes calculate the model error covariance matrix $\mathbf{Q}_i$ as part of $\mathbf{P}_i^- = \mathbf{P}_i^{p,-} + \mathbf{Q}_i$, assuming that predictability error, which is the error in the forecasts due to errors in the previous analysis or initial conditions (with covariance $\mathbf{P}_i^{p,-}$), and model error (with covariance $\mathbf{Q}_i$) are independent (Daley 1992a). One of the schemes that allows estimation of all individual elements in $\mathbf{Q}_i$ is the method of Myers and Tapley (1976). It introduces the analysis (a posteriori estimate) as a proxy for the true state and an approximation of model noise $\mathbf{w}_i^*$:

$$\mathbf{w}_i^* = \hat{\mathbf{x}}_i^+ - \hat{\mathbf{x}}_i^-, \tag{6}$$

which in fact equals the analysis increment $\mathbf{K}_i[\mathbf{y}_i - \mathbf{H}_i\hat{\mathbf{x}}_i^-]$. For an optimal Kalman filter, the covariance of this quantity is given by:

$$\begin{aligned} E[\mathbf{w}_i^*, \mathbf{w}_i^*] &= \mathbf{K}_i[\mathbf{H}_i\mathbf{P}_i^-\mathbf{H}_i^T + \mathbf{R}_i]\mathbf{K}_i^T & (7)\\ &= \mathbf{P}_i^-\mathbf{H}_i^T\mathbf{K}_i^T & (8)\\ &= \mathbf{P}_i^- - \mathbf{P}_i^+ & (9)\\ &= \mathbf{P}_i^{p,-} + \mathbf{Q}_i - \mathbf{P}_i^+, & (10) \end{aligned}$$

with $E[\mathbf{w}_i^*, \mathbf{w}_i^*] = E[(\mathbf{w}_i^* - \overline{\mathbf{w}_i^*})(\mathbf{w}_i^* - \overline{\mathbf{w}_i^*})^T]$ and $\overline{\mathbf{w}_i^*}$ the theoretical ensemble mean. The matrix $\mathbf{P}_i^+$ is the a posteriori (analysis) error covariance matrix and $\mathbf{P}_i^{p,-}$ can be calculated by propagating the previous a posteriori (or analysis) error covariance at $i-1$ by the model to time i (for a linearized model version $\mathbf{F}_{i,i-1}$: $\mathbf{P}_i^{p,-} = \mathbf{F}_{i,i-1}\mathbf{P}_{i-1}^+\mathbf{F}_{i,i-1}^T$). The practical expression to obtain matrix $\mathbf{Q}_i$ from Eq. 10 introduces a time averaging over a window of λ

time steps:

$$\hat{\mathbf{Q}}_\lambda = \frac{1}{\lambda}\sum_{i=1}^{\lambda}\left\{[\mathbf{w}_i^* - <\mathbf{w}^*>][\mathbf{w}_i^* - <\mathbf{w}^*>]^T - \left(\frac{\lambda-1}{\lambda}\right)[\mathbf{P}_i^{p,-} - \mathbf{P}_i^{+}]\right\} \tag{11}$$

An analysis increment time series (window of λ elements) is stored and centralized by its time mean $<\mathbf{w}^*>$, to eliminate the time mean bias. The time mean approach is basically to obtain a statistically viable estimate of the analysis increment covariance matrix as an approximation of the ensemble mean analysis increment covariance matrix. The advantage of this method, as compared to the computationally more attractive approach of Dee (1995), is its ability to tune the variance at each state variable individually, while simpler methods typically come up with an inflation factor to homogeneously scale the a priori covariance matrix by a constant value. The latter approach could yield adverse results, as suggested by the findings of Crow and van Loon (2006).

While in the above deduction, $\mathbf{Q}_i$ is the unknown part of $\mathbf{P}_i^- = \mathbf{P}_i^{p,-} + \mathbf{Q}_i$, we now have different missing information in $\mathbf{P}_i^-$. With the EnKF, the objective is to determine $\mathbf{P}_i^-$ based on information in the ensembles: a good approach of the error covariances would result from the propagation of the ensemble through the model dynamics. However, here the model does not simulate lateral flow, so only a part of the actual $\mathbf{P}_i^-$ is captured by the ensemble information, i.e. $\mathbf{P}_{ens,i}^-$, which is block diagonal and contains some predictability and model error (i.e. part of $\mathbf{P}_i^{p,-}$ and $\mathbf{Q}_i$ respectively) in the diagonal blocks by perturbing the initial conditions, parameters and forcings for each profile individually. The ensembles do not provide any information on the covariance between errors in variables of different profiles (off-block-diagonal), because of a model structural deficiency (no lateral flow). This is now the missing information in $\mathbf{P}_i^-$. A new matrix, $\mathbf{P}_{ad,i}^-$, will be calculated in the adaptive

scheme to adapt the ensemble-based block-diagonal $\mathbf{P}^-_{ens,i}$ to obtain the total $\mathbf{P}^-_i$:

$$\mathbf{P}^-_i = f(\mathbf{P}^-_{ens,i}, \mathbf{P}^-_{ad,i}) \tag{12}$$

The method of Myers and Tapley (1976) is slightly changed to obtain $\mathbf{P}^-_{ad,i}$ as follows:

$$\hat{\mathbf{P}}^-_{ad,\lambda} = \frac{1}{\lambda}\sum_{i=1}^{\lambda}\left\{[\mathbf{w}^*_i - <\mathbf{w}^*>][\mathbf{w}^*_i - <\mathbf{w}^*>]^T - \left(\frac{\lambda-1}{\lambda}\right)[\mathbf{P}^-_{ens,i} - \mathbf{P}^+_i]\right\}, \tag{13}$$

assuming that for $i = \lambda$: $\mathbf{P}^-_\lambda = \mathbf{P}^-_{ens,\lambda} + \mathbf{P}^-_{ad,\lambda}$, instead of $\mathbf{P}^-_\lambda = \mathbf{P}^{p,-}_\lambda + \mathbf{Q}_\lambda$ as in Eq. 11. The practical implementation is as follows: at each time step, (i) calculate the matrix $[\mathbf{w}^*_i - <\mathbf{w}^*>][\mathbf{w}^*_i - <\mathbf{w}^*>]^T$ and (ii) derive $\mathbf{P}^-_{ens,i}$ and $\mathbf{P}^+_i$ from the ensemble state members, the latter after having filtered observations with $\mathbf{P}^-_i = \mathbf{P}^-_{ens,i}$. Then, the time mean at the right hand side of Eq. 13 is calculated. In our example, the estimate of $\mathbf{P}^-_{ad,40}$ is based on increment information in the $\lambda = 40$ previous assimilation time steps for a filtering frequency of once every 2 days, from 21 December 2001 through 19 March 2002. For $\lambda \leq 20$, we found that the retrieved matrix information could not be statistically consistent, because it was too sensitive to information brought in at additional time steps.

Once $\mathbf{P}^-_{ad,\lambda}$ is found in a first period, it can be used to optimize $\mathbf{P}^-_i$ for the filtering in a next assimilation period. The variances on the diagonal of $\mathbf{P}^-_{ad,\lambda}$ are used to correct the ensemble variances on the diagonal of $\mathbf{P}^-_{ens,i}$ at each assimilation step i, i.e.:

$$[\mathbf{P}^-_i]_{kk} = [\mathbf{P}^-_{ens,i}]_{kk} + [\mathbf{P}^-_{ad,\lambda}]_{kk} \tag{14}$$

If $[\mathbf{P}^-_{ad,\lambda}]_{kk} < 0$ and the resulting $[\mathbf{P}^-_i]_{kk}$ would result in a negative value, then $[\mathbf{P}^-_i]_{kk}$ is not corrected by Eq. 14, but we choose to set $[\mathbf{P}^-_i]_{kk}/2$. Next, the covariances within

the blocks are updated for the corrected variances $[\mathbf{P}_i^-]_{kk}$ (the ensemble correlations are unaltered, the variances are updated) as follows:

$$[\mathbf{P}_i^-]_{kl} = \frac{[\mathbf{P}_{ens,i}^-]_{kl}}{\sqrt{[\mathbf{P}_{ens,i}^-]_{kk}} \cdot \sqrt{[\mathbf{P}_{ens,i}^-]_{ll}}} \cdot \sqrt{[\mathbf{P}_i^-]_{kk}} \sqrt{[\mathbf{P}_i^-]_{ll}}, \tag{15}$$

if k and l belong to the same profile. Outside the blocks, the covariances are calculated with the time-invariant correlations found in $\mathbf{P}_{ad,\lambda}^-$ and the corrected time-variant variances from Eq. 14:

$$[\mathbf{P}_i^-]_{kl} = \frac{[\mathbf{P}_{ad,\lambda}^-]_{kl}}{\sqrt{[\mathbf{P}_{ens,\lambda}^-]_{kk} + [\mathbf{P}_{ad,\lambda}^-]_{kk}} \cdot \sqrt{[\mathbf{P}_{ens,\lambda}^-]_{ll} + [\mathbf{P}_{ad,\lambda}^-]_{ll}}} \cdot \sqrt{[\mathbf{P}_i^-]_{kk}} \sqrt{[\mathbf{P}_i^-]_{ll}}, \tag{16}$$

if k and l belong to different profiles. $[\mathbf{P}_{ens,\lambda}^-]_{kk}$ is the ensemble-based a priori error variance for the kth state variable, averaged over the λ assimilation time steps. The off-block diagonal elements in $\mathbf{P}_{ad,\lambda}^-$ contain covariances containing variances of $\mathbf{P}_{ens,\lambda}^-$ (with zero off-block diagonal elements), plus $\mathbf{P}_{ad,\lambda}^-$ (the additional correction). The final $\mathbf{P}_i^-$ is not imposed on the ensemble state members along the model run, but only used for the calculation of the Kalman gain. In summary, we apply a hybrid method to fully determine $\mathbf{P}_i^-$: ensemble-based state-dependent variances are corrected by the adaptive information and are used together with (i) ensemble-based state-dependent correlations in the blocks, and (ii) time-invariant correlations outside the blocks. Unlike many EnKF studies, here, no covariance localization (Hamill et al. 2002; Reichle and Koster 2003) is applied after filling the off-block-diagonal elements in $\mathbf{P}_i^-$.

b. Innovations vs. analysis increments

The proposed method uses the information in the analysis increments ($\mathbf{K}_i[\mathbf{y}_i - \mathbf{H}_i\hat{\mathbf{x}}_i^-]$), rather than the innovations ($\mathbf{y}_i - \mathbf{H}_i\hat{\mathbf{x}}_i^-$), which are used in most other adaptive filters. A direct consequence is that all involved matrices and the resulting $\mathbf{P}_{ad,\lambda}^-$ are determined in full state space, while other adaptive filters mostly reduce the computations to the observation space. The traditional innovation method (Hollingsworth and Lönnberg 1986; Daley 1991; Houser et al. 1998) implicitly provides a way to find the a priori state (background) error correlations at observed locations and use them to cover the a larger part of state space by fitting a distance-dependent correlation function. This technique builds on the following second order innovation statistics for each pair (k, l) of observation locations:

$$E[\mathbf{r}_i^-, \mathbf{r}_i^-]_{kl} = [\mathbf{H}_i\mathbf{P}_i^-\mathbf{H}_i^T]_{kl} \quad \text{for k} \neq \text{l} \tag{17}$$

$$= [\mathbf{H}_i\mathbf{P}_i^-\mathbf{H}_i^T + \mathbf{R}_i]_{kk} \quad \text{for zero separation} \tag{18}$$

with $E[\mathbf{r}_i^-, \mathbf{r}_i^-]$ the ensemble covariance matrix at one time step i. In practice, time series of (ensemble mean) innovations $\mathbf{r}_k^-$, calculated during the EnKF application at individual locations k can be used as a proxy to estimate the ensemble innovation covariance. The centralized second order statistic is calculated to remove the time mean discrepancy between forecasts and observations. Because in our study, a reasonable time-invariant estimate of the diagonal $\mathbf{R}$-matrix is available (i.e. $[\mathbf{R}]_{kk} = (0.022\ \mathrm{m^3m^{-3}})^2$ and $[\mathbf{R}]_{kl} = 0\ (\mathrm{m^3m^{-3}})^2$), the a priori error correlations ρ_{kl} can be estimated by:

$$\rho_{kl} = \frac{E[[\mathbf{r}^-]_k, [\mathbf{r}^-]_l] - [\mathbf{R}]_{kl}}{\sqrt{(E[[\mathbf{r}^-]_k, [\mathbf{r}^-]_k] - [\mathbf{R}]_{kk})(E[[\mathbf{r}^-]_l, [\mathbf{r}^-]_l] - [\mathbf{R}]_{ll})}} \tag{19}$$

with $E[[\mathbf{r}^-]_k, [\mathbf{r}^-]_l] = E[([\mathbf{r}^-]_k - < [\mathbf{r}^-]_k >)([\mathbf{r}^-]_l - < [\mathbf{r}^-]_l >)^T]$ calculated over time and $< [\mathbf{r}^-]_k >$ the time mean innovation for the k th element of vector $\mathbf{r}^-$. For each pair of elements this correlation can then be plotted as function of the distance and an empirical homogeneous (in space) function can then be fitted.

In this study, one could easily change the above procedure, or any other innovation-based adaptive method, to find the off-block-diagonal correlation values of $\mathbf{P}^-_{ad,\lambda}$ in the observation space to fill in the individual entries in $\mathbf{P}^-_i$ in the state space, as well as to find the variance correction, like in our adapted Myers and Tapley (1976) method. However, that would only be possible in the case of this study, where $\mathbf{H}_i$ only contains 0 and 1 and the transfer of information from the observation to the state space would be well determined (leaving unobserved entries in $\mathbf{P}^-_i$ with some default or zero correlation). For more general and complex observation operators, the method of Myers and Tapley (1976) is more attractive than other adaptive filters to find $\mathbf{P}^-_{ad,\lambda}$ in the state space.

5. Results

a. Innovation analysis: correlation structure

Before implementing the adaptive Myers and Tapley (1976)-based method described above, it was tested if a time-invariant horizontal error correlation structure could be defined as function of separation distance between profiles, based on the information in the innovation covariances, as in the traditional innovation method.

For a 2-weekly assimilation of complete profiles from 2 October 2001 through 19 March

2002, the resulting 13-element time series of (ensemble mean) innovations were used to approximate a time-invariant ensemble innovation covariance. The deduced background error correlations in Figure 3 show that it is impossible to fit an empirical homogeneous isotropic correlation function to the calculated forecast error correlations at any observation depth. Similar results were found for scenarios with more frequent assimilation (and hence more elements in the innovation time series).

The CLM does not simulate horizontal water flows and is applied for each profile individually, while studies in the OPE^3 field have revealed pathways for active preferential horizontal flow (Gish et al. 2002). It could be expected that profiles situated along these pathways may therefore show correlated model structural errors and hence correlated innovations. However, there were too few reported actively connected probes with available observations (and hence innovations) at corresponding soil depths to quantitatively analyze this hypothesis. Furthermore, the innovation correlation due to correlation in model structural errors may have been reduced by error compensation during the individual profile calibration procedure.

The lack in spatial error structure in Figure 3 is probably because (i) the error correlation structure changes with direction (i.e. anisotropy), (ii) the different profiles were calibrated independently and (iii) this is a small field scale analysis, where the local highly variable soil conditions (and their potential parameterization errors) largely determine a close to random spatial soil moisture structure (and hence the derived innovation statistics). Note that at the regional scale, one might find spatial correlation structures controlled by forcing error patterns.

Figure 3 also shows that error correlations are higher and less variable in shallow layers than in deeper profile layers. Similarly, the analysis of the spatial structure in the soil

moisture observations showed that the correlation (and correlation length) was higher in shallow layers (De Lannoy et al. 2006b). This suggests that for highly correlated time series of observed soil moisture, also highly correlated innovation time series are found. This is reasonable, because the calibrated model deficiencies will be similar for similar observed soil moisture time series. Basically, the innovation time sequences will contain sequences of short time scale biases which might be correlated. When a sequential bias estimation (De Lannoy et al. 2007b) was included and the bias was removed from the innovations at each assimilation step, before the innovation time series was centralized by its time mean (longer time scale 'bias'), the correlations in the top layers indeed dropped down (results not shown).

To study the correlation stationarity in time, daily assimilation of complete profiles was performed over the same period (2 October 2001 through 19 March 2002, now resulting in 167-element time series). Through batch processing, the correlation was calculated over 5 (and 10) windows of 33 (and 16) hourly time steps. A random selection of soil moisture state variables pairs at identical depths (10 cm in Figure 4) shows that the retrieved error correlation values are non-stationary at the high resolution time scale (e.g. 10 windows), but change relatively little when smoothed over less windows.

b. Adaptive filter: spatial effects

For the main experiment, the combined state estimation (updating) and uncertainty estimation was performed during the period from 2 October 2001 through 21 December 2001 with the adapted Myers and Tapley (1976) method. All 36 complete observed soil moisture profiles were assimilated every 2 days (40 time steps) to calculate missing off-

block-diagonal information and to obtain a correction to the variances in the a priori error covariance. Thereafter, from 23 December 2001 until 19 March 2002, only a single profile was assimilated using the novel error covariance information to enhance the soil moisture estimates at all 36 simulated soil moisture profiles in the field. During this period the ensemble-based $\mathbf{P}^{-}_{ens,i}$ was updated with time-invariant information in $\mathbf{P}^{-}_{ad,\lambda}$ as described in Eqs. 14 through 16.

Figure 5 shows the a priori state error correlation matrix on 23 December 2001, prior to and after the adaptive correction. It illustrates how the ensemble-based block-diagonal (i.e. within a profile) correlations are kept unaltered and how the off-block-diagonal (i.e. between profiles) error correlation elements are partially filled in. For a random zoom example, where error correlations between state variables corresponding to profile AL2 and BM4 are shown, only 6 rows (state variables corresponding to the 6 observation layers in M-probes) and 7 columns (7 observation layers in L-probes) are filled in. The more observational layers are available for a pair of profiles, the more error correlation information can be retrieved. Because the model layers at 2.5, 5 and 20 cm are never observed by the OPE^3-probes, there will never be information on error statistics at these layers. Again no spatial structure could be found in the retrieved error correlations at any soil layer (not shown). Similar to the above findings, the determined error correlation was higher in the top soil layer, but much smaller than what was found through the innovation method.

Figure 6 shows for 2 individual soil depths at 10 and 80 cm, the temporal evolution of the Root spatial Mean Square Error (RMSE) between observations and ensemble mean forecasts without assimilation as well as analyses with assimilation of complete observed soil moisture profile data. During the first period (2 October 2001-21 December 2001) all available profile

information is assimilated, causing significant drops in RMSE. However, the observational information is not taken up completely by the model at all profiles, mainly due to forecast bias (sometimes the model and observation climatologies are different, De Lannoy et al. (2007b)), causing a rapid increase of RMSE after each assimilation event. In the second period, only 1 selected profile (here, DL1) is assimilated, either without or with inclusion of the estimated off-block-diagonal a priori error covariance elements. The further reduced RMSE when including the off-block-diagonal information shows that there is a benefit in spreading information in the horizontal space.

As an illustration, the spatial soil moisture fields in Figure 7 are generated for 13 March 2002 as a weighted average of spatial interpolations (only for visualization purposes) by (i) a multivariate regression between the soil moisture and the terrain characteristics (texture, elevation, topographic index) at the observing locations, and (ii) kriging with an exponential variogram with parameters as observed in De Lannoy et al. (2006b). The observations and simulations show a discrepancy in the spatial soil moisture field: a part is due to model bias, another part is random error. When assimilating the complete observed profile at DL1 with the regular EnKF, no other profile is updated and the spatial view is dominated by the ensemble mean integration without assimilation. When the estimated off-block-diagonal a priori error covariance elements are included, ADEnKF of DL1 observations updates all profiles in the field. The full circles in the figure show regions where the ADEnKF obviously brings the simulated field close to the observed one, while only assimilating data at 1 location in the field and using time-invariant correlation information off the diagonal blocks in $\mathbf{P}_i^-$. The dashed circle in field C shows a region where a major dry forecast bias cannot be dealt with by a state filter and where the adaptive filter only has a minor effect, because it is mainly

designed to treat random error and probably also because of the time-invariant correlation structure. There is an attempt to increase the soil moisture in field C by the ADEnKF, but it is mainly effective only at 1 location (CL2).

Figure 8 summarizes the effect on the spatial soil moisture estimation, when assimilating observations from each individual profile only, during a validation period from 23 December 2001 through 24 March 2002. The time averaged spatial RMSE is shown for the case without and with inclusion of $\mathbf{P}^{-}_{ad,40}$ information. The RMSE is generally smaller with inclusion of off-block-diagonal a priori error covariance elements, even though the off-block-diagonal structure was determined over the period preceding the validation period and kept time-invariant (and hence suboptimal) over the validation period. A maximum decrease in spatial RMSE by 22% was achieved at 180 cm by assimilating either DM2 or DM3 profile data, a maximum RMSE decrease of 13% was found at 80 cm for the assimilation of AM1 data and at 10 cm a 9% RMSE decrease resulted from assimilation of BM2 data. At the other observing layers, the maximum RMSE decrease was 11% (30 cm), 13% (50 and 80 cm), 17% (120 cm) and 7% (150 cm), by complete profile assimilation. For this study, the maximum decrease in spatial RMSE, averaged over all field profile layers was 11%, when assimilating probe AM1. It turns out that AM1 is the most time stable profile in the OPE^3 field (De Lannoy et al. 2007a), i.e. its measured soil moisture time series best resembles the spatial mean soil moisture behaviour in time at most soil depths. It should also be noted that validation of the performance in a single layer sometimes shows further reduced spatial RMSE values, when only assimilating a single observation in that single layer (not shown). This is because when updating the soil moisture in many layers simultaneously, the analysis profile may not reflect the preferred model balance and seemingly contradictory (to the model) updates for

the different layers may be cancelled out quickly over time after the analysis. The update in a single layer often persists longer in time, when neighbouring layers are adapting to the update in that single layer, rather than possibly competing against this update with updates at their own respective depths. It is evident that, when evaluating the total field profile performance (as opposed to the previous discussion on the effect in a single layer), assimilation with probes having more observational depths (6 for M-probes, 7 for L-probes) generally results in the best improvements in RMSE. However, it is important to notice how H-probes with only 3 observation layers are able to enhance the spatial RMSE at unobserved (e.g. 180 cm) layers.

In some cases (here for 80 cm depth), the RMSE is increased, either due to a bad propagation of information over the vertical layers or to inaccurate determination of the between-profiles error correlation. In an exceptional case, the assimilation of a single profile did not improve the soil moisture field estimate in any layer: through assimilation of probe BL2, the spatial RMSE was increased outside the range of the plot for almost all soil layers when including the spatial error correlation information. The problem here was a collapsed ensemble spread (and the time-invariant adaptive correction was inappropriate to solve this problem) at some layers, where a model-defined minimal soil moisture was reached, causing adverse effects in other layers at distant profiles.

The spatial RMSE covers both the discrepancy in the spatial soil moisture field and the spatial structure of the soil moisture field. The last 3 columns in Table 1 summarize the root time mean square difference between the observed and modelled spatial mean soil moisture field (over the validation period 23 December 2001 - 24 March 2002) for the cases (i) without assimilation, (ii) with regular EnKF assimilation of a single profile (here, DL1)

and (iii) with assimilation of the same single profile by the proposed adaptive EnKF. When only state updating is performed (EnKF), the RMSE is slightly reduced, only because of the improving effect on the DL1 soil moisture time series within the spatial mean. Clearly, when including adaptive information, the error in the spatial mean soil moisture is largely reduced by assimilating a single profile. Only at 150 cm (for this case), a minor adverse effect can be observed caused by a bad vertical interaction between layers. Maybe this could be solved by using the adaptive filter to correct the ensemble-based error correlations within the profile. The first 3 columns in Table 1 give the time mean of the root mean square difference in semivariances γ between the observed (γ_o) and modelled (γ_m) spatial soil moisture fields over the same period of length T:

$$\frac{1}{\mathrm{T}}\sum_{i=1}^{\mathrm{T}}\sqrt{\frac{1}{\mathrm{L}}\sum_{l=1}^{\mathrm{L}}(\gamma_{o,l}-\gamma_{m,l})_i^2} \tag{20}$$

The spatial lag in the calculation of γ_o and γ_m was set at 20 m (to allow some averaging over pairs of observations in each spatial lag-interval) for a maximum interdistance of 420 m (beyond half of the field length, semivariance values would be calculated for too few sample pairs): the mean square difference in semivariances was thus calculated over a maximum number of 21 spatial lags, i.e. L = 21. The Table 1 shows that, through simple EnKF, the spatial pattern was slightly improved, again because of the enhanced representation of soil moisture at the assimilation location only. With inclusion of adaptive information, the differences in semivariances were further reduced. Consequently, assimilation of a single profile (here, DL1) with an adaptive EnKF is able to partially rectify both incorrectly simulated spatial soil moisture mean and variance.

6. Discussion and perspectives

In this paper, the missing error correlations and an error covariance correction are retrieved only over a single time window, and to be used in a period succeeding the period where it was actually calculated. A more frequent batch-processing could be suggested to obtain more time-variant adaptive information. Definitely for large scale applications, where the variability in the precipitation errors may alter the error correlations over time, a frequent update of the missing error correlations would be beneficial. However, there is a trade-off between an increased time variability and the statistical consistency of the error covariance estimates as found in Eq. (13): some time averaging is needed to obtain viable covariance estimates. Future research could investigate how the training period can be limited by exploring the ensemble information in the innovations at each time step, instead of studying the ensemble mean information in time.

The analysis results can be further improved by adding a bias filter to the state estimation (De Lannoy et al. 2007b). However, since the bias error covariance is often linked to the (adapted) a priori error covariance, there is more research needed to assure both proper interactions between these error covariances and optimality of the filter. Note that the Myers and Tapley (1976) method removes some bias from the increments to find $\mathbf{P}^{-}_{ad,\lambda}$ in Eq. 13, but the state estimates remain biased. It should also be recognized that the a priori error correlations should only represent information about the random error structure: long range a priori state error correlations derived from innovations (or analysis increments) may sometimes originate from spatial bias fields in either the observations or forecasts.

This approach to determine the between-profiles error covariance can be of practical im-

portance to propagate satellite observations from observed locations to temporarily masked or unobserved regions (e.g. clouds in visual products, hydrometeors for high-frequency microwave products), or propagation of in situ profile information to locations where sensors are temporarily out of service. In satellite data assimilation schemes, the modelled grid cells are typically covered by the observations (after reprojecting the data). This allows to update all the a priori error correlations at each time observations are available and use them, whenever part of the observations falls out and information needs to be propagated to the non-observed pixels. Furthermore, training of the correlations during one satellite pass (covering a limited banded area) allows observations in the overlapping area of the next pass to be propagated all over the area covered by previous (shifted) pass, if some time-invariance in error correlations could be assumed. However, for large scale applications, one generally assumes that the a priori error correlation would be limited (e.g. different dominant forcing errors) and not worth the increased numerical expenses to deal with a more complex $\mathbf{P}_i^-$. As a very simple approximation, the actual (mostly observed) state correlation length is often used to limit the a priori state error correlations at large distances, even though one should realize that the actual state and its estimation errors (and hence their respective correlations lengths) are in fact independent.

A problem is that with a finer model resolution, more observations are needed to retrieve the correct error covariances. If a soil moisture profile is available every 20 m (assume it representative for a 10×10 m^2 area), and the model simulates at a 10 m resolution, then the correlation beween the forecast errors of 2 neighbouring grid cells at 10 m apart will not be retrieved by the proposed technique, while the correlation between forecast errors in 2 observed grid cells at 20 m apart can be found. One limitation of the proposed method is

the need to determine the adaptive information based on as many observations as available, before it can be used with a restricted dataset.

In general, the proposed technique is expected to advance the optimal use of satellite and other data to fill in unobserved areas and to overcome the shortcomings of completely time-invariant homogeneous (in space) correlation matrices as in optimal interpolation techniques. Typical empirical distance-dependent correlation functions may give a false impression of smooth correlation structures in the analyses. It could be advised to review assumptions on empirical correlation functions for hydrological variables at different scales by dynamical error correlation insights obtained from adaptive filters.

7. Summary

This study explores the potential of updating the full 3D system state when assimilating only single soil moisture profile data in the independent column-based CLM2.0 with an adaptive ensemble Kalman filter (ADEnKF). Assimilation of individual in situ soil moisture profiles from the small scale OPE^3 field with an optimized knowledge of a priori error covariances resulted in a partial rectification of the inaccurately simulated spatial mean and variability in soil moisture fields at different soil depths.

Because no simple spatial error correlation structure could be found based on the traditional innovation method, it was necessary to find all elements of the a priori error covariance matrix individually. Ensembles were used to obtain (i) a first guess of each state variable uncertainty, i.e. the magnitude of error variance and (ii) an estimate of the error correlation between state variables within each individual profile. Then, during a first period, a

time-invariant variance correction and the between-profiles error correlations were sought based on an adaptive method of Myers and Tapley (1976). This allowed adjustment of the time-variable ensemble-based state error covariance matrix before each filtering time step in the successive period.

Even though only one time-invariant estimate was used to correct the a priori error covariances, the results show improvements up to between 7 and 22% reduction in spatial RMSE in a single soil layers, when assimilating only a single complete profile (higher percentage reductions can be achieved by assimilating only in the single validation layer, which avoids contradictory updates over the different soil layers, not shown). Even assimilation of a single profile with only a few observed layers resulted in enhanced soil moisture field estimates at most depths. With knowledge of the off-block-diagonal a priori error covariances (ADEnKF), any observation in a single layer will affect both the total profile where it is assimilated in and all layers of surrounding profiles, while a regular EnKF with the CLM2.0 only allows updating of a single profile.

This study shows an example of retrieving the missing error correlations and an error covariance correction during a first time window, and using it in a period succeeding the period where it was actually calculated. A more frequent batch-processing would result in more time-variant adaptive information and the information in the ensembles may be further explored to limit the training periods, while keeping the estimates statistically consistent.

The determination of between-profiles error covariance can be important to propagate satellite observations from observed locations to temporarily unobserved or masked regions, or propagation of in situ profile information to locations where sensors are temporarily out of service. The limitation of this method is the need to determine the adaptive information

based on as many observations as available, before it can be used with a restricted dataset.

Acknowledgments.

The first author is a postdoctoral research fellow of the Research Foundation Flanders (FWO). We thank the USDA Beltsville Agricultural Research Center/Agricultural Research Service for their data. The reviewers are thanked for their constructive comments.

REFERENCES

Anderson, J. L., 2007: An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus*, **59A**, 210–224.

Crow, W., 2003: Correcting land surface model predictions for the impact of temporally sparce rainfall rate measurements using an ensemble Kalman filter and surface brightness temperature observations. *Journal of Hydrometeorology*, **4**, 960–973.

Crow, W. T. and E. van Loon, 2006: Impact of incorrect model error assessment on the sequential assimilation of remotely sensed surface soil moisture. *Journal of Hydrometeorology*, **7**, 421–432.

Dai, Y., et al., 2003: The Common Land Model. *Bulletin of the American Meteorological Society*, **84**, 1013–1023.

Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, 455 pp. pp.

Daley, R., 1992a: Estimating model-error covariances for application to atmospheric data assimilation. *Monthly Weather Review*, **120 (8)**, 1735–1746.

Daley, R., 1992b: Forecast-error statistics for homogeneous and inhomogeneous observa tion networks. *Monthly Weather Review*, **120 (4)**, 627–643.

De Lannoy, G. J. M., P. R. Houser, V. R. N. Pauwels, and N. E. C. Verhoest, 2006a:

Assessment of model uncertainty for soil moisture through ensemble verification. *Journal of Geophysical Research*, **111 (D10)**, D10 101.1–18, doi:10.1029/2005JD006367.

De Lannoy, G. J. M., P. R. Houser, N. E. C. Verhoest, V. R. N. Pauwels, and T. Gish, 2007a: Upscaling of point soil moisture observations to field averages at the OPE^3 site. *Journal of Hydrology*, 1–11, doi:10.1016/j.jhydrol.2007.06.004.

De Lannoy, G. J. M., R. H. Reichle, P. Houser, V. R. N. Pauwels, and N. E. C. Verhoest, 2007b: Correcting for forecast bias in soil moisture assimilation with the ensemble kalman filter. *Water Resources Research*, **43 (6)**, W09 410.1–14, doi:10.1029/2006WR00544.

De Lannoy, G. J. M., N. E. C. Verhoest, P. R. Houser, T. Gish, and M. Van Meirvenne, 2006b: Spatial and temporal characteristics of soil moisture in an intensively monitored agricultural field (OPE^3). *Journal of Hydrology*, **331 (3-4)**, 719–730.

Dee, D. P., 1995: On-line estimation of error covariance parameters for atmospheric data assimilation. *Monthly Weather Review*, **123**, 1128–1145.

Dee, D. P. and A. M. da Silva, 1999: Maximum-likelihood estimation of forecast and observation error covariance parameters. Part I: Methodology. *Monthly Weather Review*, **127**, 1822–1834.

Drécourt, J.-P., H. Madsen, and D. Rosbjerg, 2006: Joint calibration and uncertainty analysis of a groundwater model coupled with a Kalman filter. *Advances in Water Resources*, **29 (5)**, 719–734.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model

using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, **99 (C5)**, 10 143–10 162.

Gish, T., W. Dulaney, K.-J. S. Kung, C. Daughtry, J. Doolittle, and P. Miller, 2002: Evaluating use of ground-penetrating radar for identifying subsurface flow pathways. *Soil Science Society of America Journal*, **66 (5)**, 1620–1629.

Hamill, T., C. Snyder, and R. Morss, 2002: Analysis-error statistics of a quasi-geostrophic model using three-dimensional variational assimilation. *Monthly Weather Review*, **130**, 2777–2790.

Hollingsworth, A. and P. Lönnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*, **38A**, 111–136.

Houser, P. R., W. J. Shuttleworth, J. S. Famiglietti, H. V. Gupta, K. H. Syed, and D. C. Goodrich, 1998: Integration of soil moisture remote sensing and hydrologic modeling using data assimilation. *Water Resources Research*, **34 (12)**, 3405–3420.

Huang, H.-C. and N. Cressie, 1996: Spatio-temporal prediction of snow water equivalent using the kalman filter. *Computational Statistics & Data Analysis*, **22**, 159–175.

Jazwinski, A., 1969: Adaptive filtering. *Automatica*, **5**, 475–485.

Kailath, T., 1968: An innovation approach to least-squares estimation. Part I: Linear filtering in additive white noise. *IEEE-Transactions on Automatic Control*, **AC-13**, 646–655.

Kumar, P., 1999: A multiple scale state-space model of characterizing subgrid scale variabil-

ity of near-surface soil moisture. *IEEE Transactions on Geoscience and Remote Sensing*, **37 (1)**, 182–197.

Maybeck, P. S., 1982: *Stochastic Models, Estimation, and Control. Volume 2*, Mathematics in science and engineering, Vol. 423. Academic Press.

McLaughlin, D., 2002: An integrated approach to hydrologic data assimilation: interpolation, smoothing, and filtering. *Advances in Water Resources*, **25**, 1275–1286.

Mehra, R., 1970: On the identification of variances and adaptive Kalman filtering. *IEEE Transactions on Automatic Control*, **April**, 175–184.

Mehra, R., 1972: Approaches to adaptive filtering. *IEEE Transactions on Automatic Control*, **17 (5)**, 693–698.

Mitchell, H. L. and P. L. Houtekamer, 2000: An adaptive ensemble Kalman filter. *Monthly Weather Review*, **128**, 416–433.

Myers, K. and B. Tapley, 1976: Adaptive sequential estimation with unknown noise statistics. *IEEE Transactions on Automatic Control*, **21 (4)**, 520–523.

Parada, L. M. and X. Liang, 2004: Optimal multiscale Kalman filter assimilation of near-surface soil moisture into land surface models. *Journal of Geophysical Research*, **109**, D24 109.1–D24 109.21.

Reichle, R. H., W. Crow, and C. Keppenne, 2008: An adaptive ensemble Kalman filter for soil moisture data assimilation. *Water Resources Research*, **XX**, XXX – XXX, accepted.

Reichle, R. H. and R. Koster, 2003: Assessing the impact of horizontal error correlations in background fields on soil moisture estimation. *Journal of Hydrometeorology*, **4**, 1229–1242.

Reichle, R. H., D. B. McLaughlin, and D. Entekhabi, 2002a: Hydrologic data assimilation with the ensemble Kalman filter. *Monthly Weather Review*, **130 (1)**, 103–114.

Reichle, R. H., J. P. Walker, P. R. Houser, and R. D. Koster, 2002b: Extended versus ensemble Kalman filtering for land data assimilation. *Journal of Hydrometeorology*, **3 (6)**, 728–740.

Romanowicz, R. J., P. C. Young, and K. J. Beven, 2006: Data assimilation and adaptive forecasting of water levels in the river Severn catchment, United Kingdom. *Water Resources Research*, **42**, W06 407.1–W06 407.12.

Stauffer, D. R. and N. L. Seaman, 1990: Use of four-dimensional data assimilation in a limited-area mesoscale model. Part I: Experiments with synoptic-scale data. *Monthly Weather Review*, **118 (6)**, 1250–1277.

van Geer, F. C., C. B. te Stroet, and Z. Yangxiao, 1991: Using Kalman filtering to improve and quantify the uncertainty of numerical groundwater simulations. *Water Resources Research*, **27**, 1987–1994.

Walker, J. P. and P. R. Houser, 2004: Requirements of a global near-surface soil moisture satellite mission: Accuracy, repeat time, and spatial resolution. *Advances in Water Resources*, **27 (8)**, 785–801.

Zhang, S.-W., C.-J. Qiu, and A. Xu, 2004: Estimating soil water contents from soil temper-

ature measurements by using an adaptive Kalman filter. *Journal of Applied Meteorology*, **43**, 379–389.

List of Tables

List of Figures

List of Figures

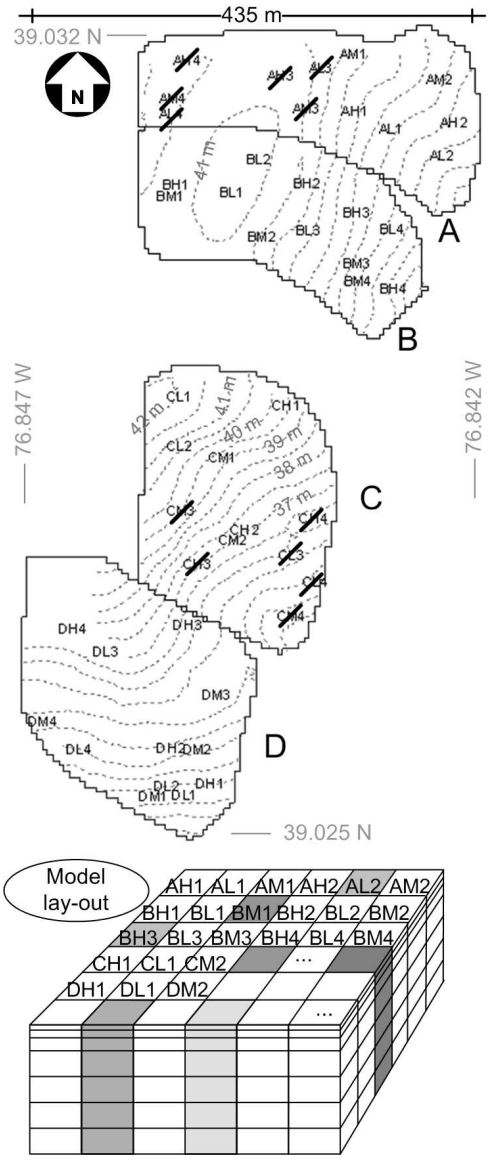

FIG. 1. Digital elevation model (dashed contour lines, 0.5 m interval) with indication of the boundary coordinates and location of soil moisture probes (3 digits) in the OPE^3 field with 4 sub-watersheds A through D. Defective probes during the study period are crossed out. The bottom sketch shows how the system is modelled as 36 individual, unconnected, soil profiles with multiple soil layers.

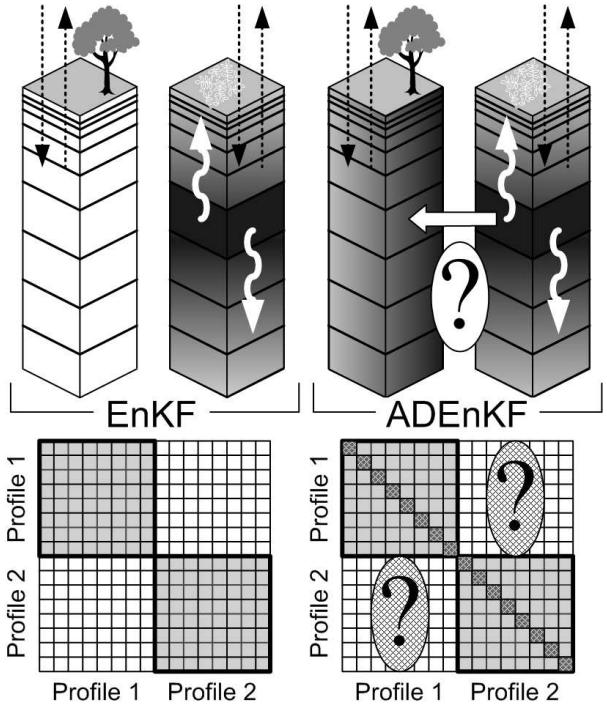

FIG. 2. Schematic of EnKF and ADEnKF functionality with the CLM, which only simulates vertical fluxes (black dashed arrows). The white arrows represent the information transfer from the observation in one particular layer to model states in other soil layers of the same profile (EnKF and ADEnKF) and neighbouring profiles (ADEnKF only) by the a priori state error covariance statistics $\mathbf{P}_i^-$ in the filter. The bottom matrices represent $\mathbf{P}_i^-$ for 2 profiles and with multiple state variables per profile. For the EnKF, $\mathbf{P}_i^-$ is block-diagonal; for the ADEnKF, it is tried to (i) optimize the a priori error variances on the diagonal of $\mathbf{P}_i^-$ and (ii) retrieve the off-block-diagonal error correlations between the different profiles, schematically indicated by crosshatched matrix parts.

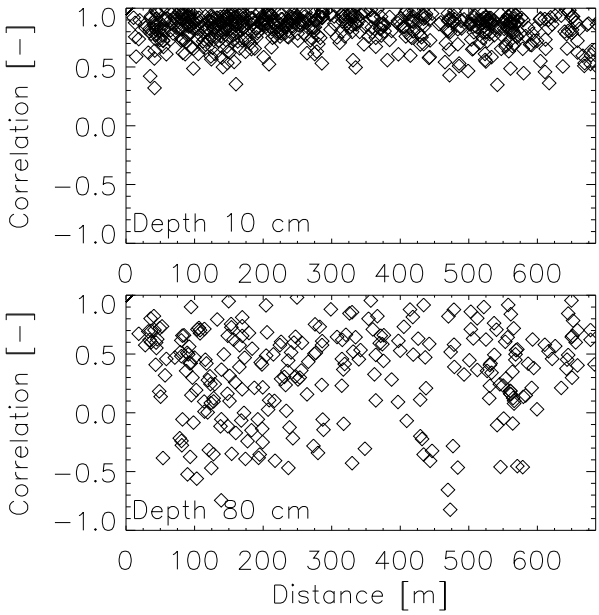

FIG. 3. A priori horizontal error correlation as function of separation distance, calculated by the traditional innovation method at 10 cm and 80 cm depth with 13 innovations obtained after assimilation of complete profiles every 2 weeks from 2 October 2001 through 19 March 2002.

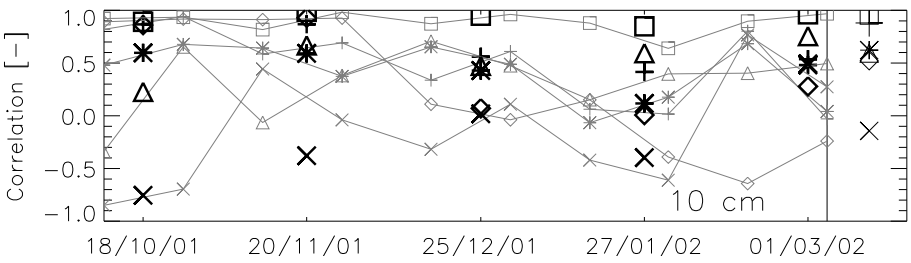

FIG. 4. Calculated error correlation at 10 cm depth for some randomly selected pairs: + (DH1,A21), ∗ (AM1,BL1), ⋄ (AH2,DH3), △ (BH2,BL2), □ (BL3,CL1), × (BM4,DM2). Small gray symbols are calculated over 10 windows of 16 hourly time steps, thick black symbols over 5 windows of 33 time steps. The symbols on the right are calculated over the total time series of 167 hourly time steps.

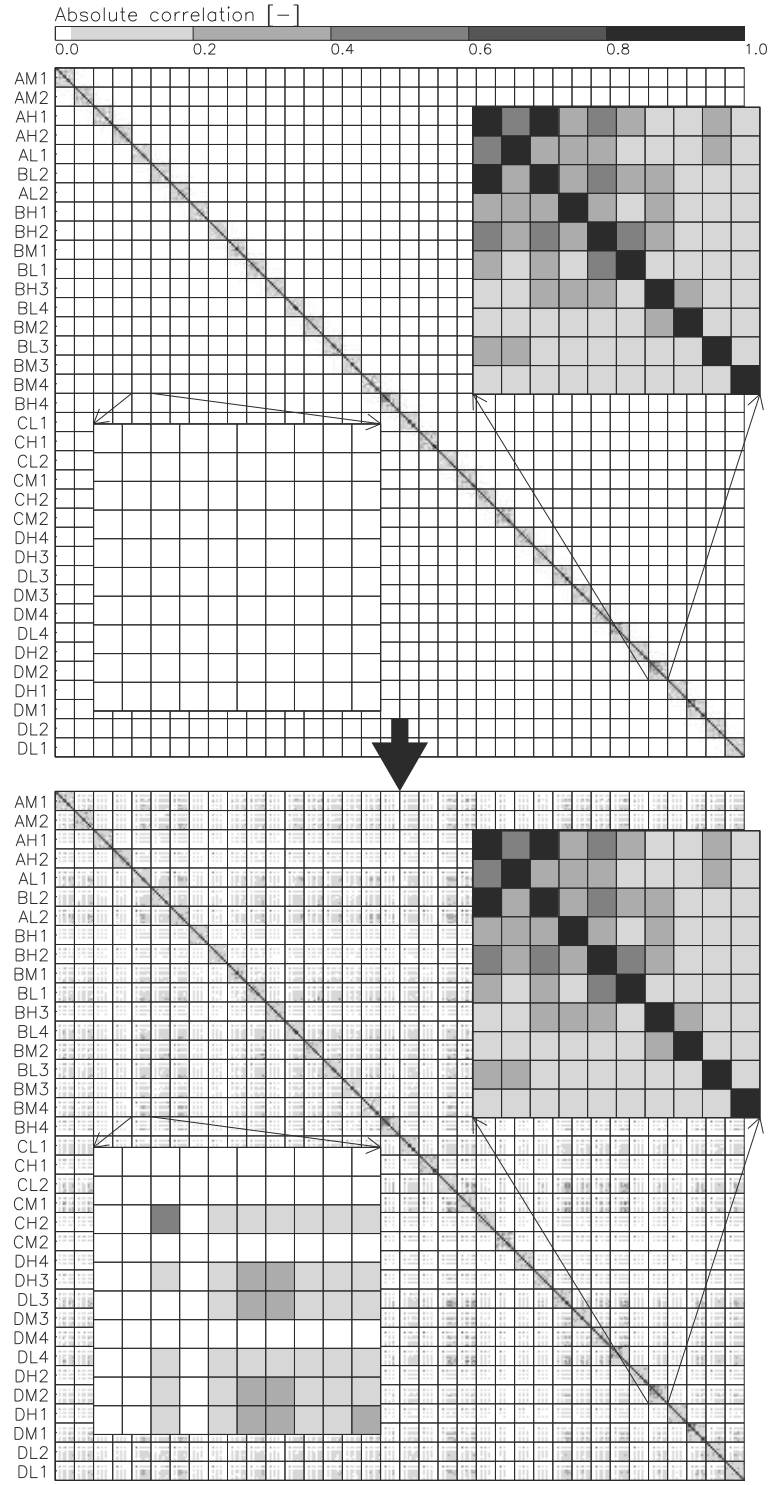

FIG. 5. A priori error correlation matrix (upper) prior to and (lower) after adaptive correction at 23 December 2001. Within each profile 10 soil moisture state variables (2.5, 5, 10, 20, 30, 50, 80, 120, 150 and 180 cm) are present. The absolute correlation values are shown.

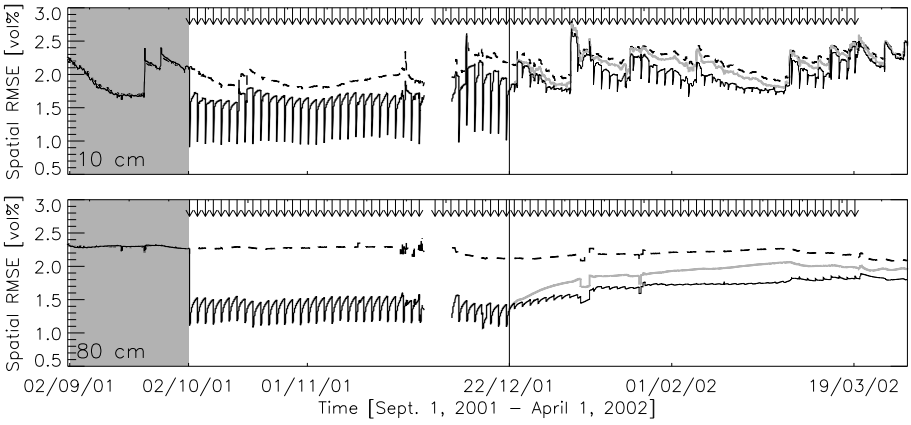

FIG. 6. Spatial RMSE in soil moisture over all available sensors at 2 different depths. The arrows indicate the assimilation events: from 2 October 2001 through 21 December 2001 all profile information is assimilated, thereafter through 19 March 2002 only profile data from probe DL1 are assimilated. The gray zone shows the model calibration period. The dashed line is for the ensemble mean integration without any assimilation, the gray line and the full black line are assimilation runs, respectively without and with inclusion of off-block-diagonal information extracted from the innovations. The gap is caused by missing observations.

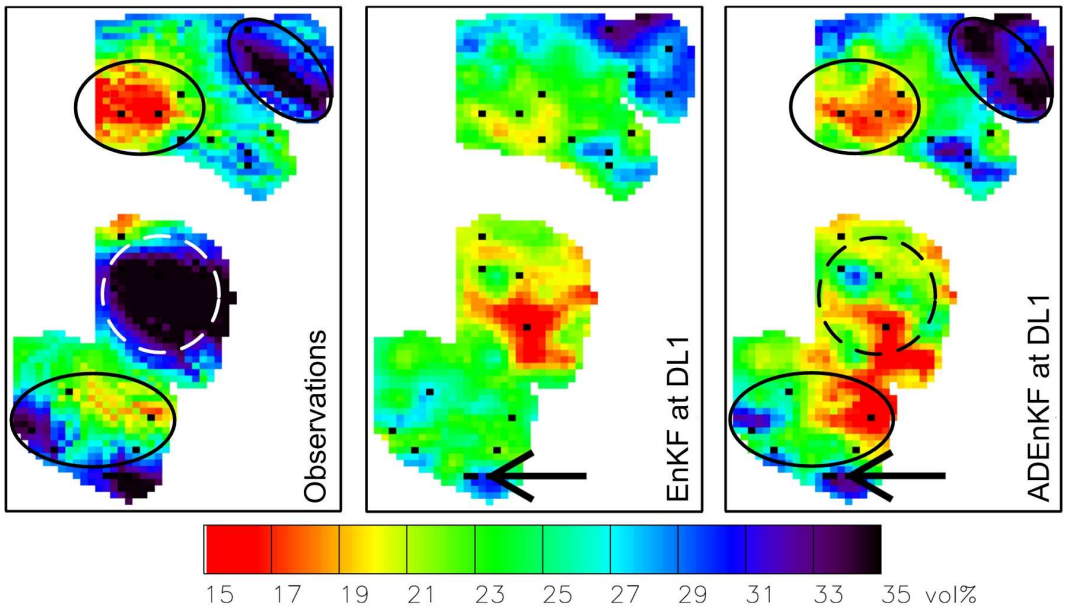

FIG. 7. Spatial soil moisture fields at 50 cm depth on 13 March 2002, obtained by spatial interpolation of the (left) observed values, (middle) ensemble mean forecasts with EnKF assimilation of the complete profile information at DL1 and (right) ensemble mean analyses obtained by ADEnKF assimilation of the complete profile information at DL1, with propagation of the update over the whole field. The arrows indicate the assimilation location, the full ellipses show regions where the ADEnKF has a clear improving impact, the dashed circle shows a region where the model (middle) largely underestimates soil moisture and the ADEnKF (right) only has a very limited impact. The black dots are the actual observed and simulated locations at 50 cm depth.

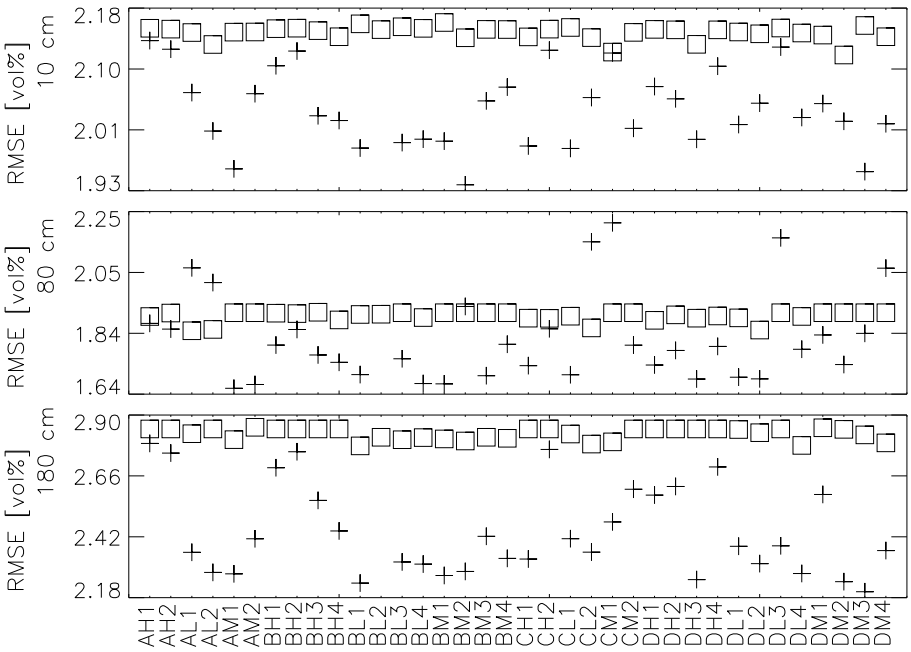

FIG. 8. Spatial RMSE in soil moisture over all available sensors at 3 different depths, averaged over a validation period of 23 December 2001 through 24 March 2002, covering the period where only a single profile (x-axis) is assimilated. The symbols □ and + are respectively without and with inclusion of off-block-diagonal information extracted from the analysis increments.

List of Tables

TABLE 1. Time averaged RMS difference between observed and modeled soil moisture spatial semivariance (DSsv) [vol%^{2}] and spatial mean (DSm) [vol%], over a validation period from 23 December 2001 though 24 March 2002. Ens stands for the ensemble mean integration without any assimilation, EnKF and ADEnKF are for 2-daily ensemble Kalman filtering of the single DL1 profile over the period 23 December 2001 thourgh 19 March 2002, without and with inclusion of adaptive information, respectively.

	DSsv [vol%^{2}]			DSm [vol%]		
Depth	Ens	EnKF	ADEnKF	Ens	EnKF	ADEnKF
10 cm	14.99	14.22	11.25	3.54	3.41	3.15
30 cm	20.61	19.83	18.34	1.57	1.24	0.64
50 cm	89.97	84.99	77.90	4.44	3.89	2.97
80 cm	34.57	26.29	23.59	3.02	2.35	0.63
120 cm	80.47	69.59	56.57	4.72	4.09	1.65
150 cm	86.59	70.72	59.62	2.13	1.61	1.76
180 cm	71.33	70.06	56.74	8.23	6.99	3.47